64-5

ELECTRONICS PERSONNEL RESEARCH GROUP

Department of Psychology

UNIVERSITY OF SOUTHERN CALIFORNIA

Technical Report No. 34

THE RELATIVE EFFICIENCY OF DIFFERENT COMBINATIONS OF PROMPTING AND CONFIRMATION FOR LEARNING A BOOLEAN ALGEBRA PROGRAM

February 1962

Project Designation NR 153-093

Contract Nonr-228(22)

Joseph W. Rigney

Irving J. Budnoff

<u>Prepared for</u>

Personnel and Training Branch
Psychological Sciences Division
Office of Naval Research

DEPARTMENT OF PSYCHOLOGY

UNIVERSITY OF SOUTHERN CALIFORNIA

ACKNOWLEDGMENTS

The collection of data for this study was made possible through the cooperation of Mr. Norman R. Dilley, Head of the Mathematics Department, Huntington Beach High School, who facilitated arrangements with the district and school principals, and who permitted the use of his mathematics class as subjects.

Dr. Edward B. Fry, Associate Professor, Loyola University, was a consultant during the initial planning of the study.

The study was sponsored by Dr. Glenn L. Bryan and Mr. John Nagay of the Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research. Their continued support and encouragement are appreciated.

ABSTRACT

This study investigated the effects of prompting (P) and confirmation (C), and of three combinations (PP, PC, and CC) of these in the learning of a Boolean algebra program.

One hundred and five high school mathematics students, classified into three intelligence levels, served as subjects. Each level was then divided into three random experimental groups (PP, PC, and CC). Subjects in each experimental group worked through each lesson twice in the learning situation. Group PP worked through the lessons both times under the prompting condition; Group PC had prompting first and confirmation second; Group CC had confirmation both times. For the PC and CC groups, the second trial served as an acquisition test. Repeated-item and transfer tests were given at the end of 7- and 21-day intervals to all groups.

It was hypothesized that: (1) the PC combination would be superior to the CC combination for acquisition of associations, and to the PP combination for retention of associations; (2) the PC and CC combinations would produce higher transfer scores than the PP combination; and (3) the brightest students would be less affected by these combinations of experimental variables than the other students.

None of the above hypotheses was confirmed by analysis of the main effects of the experimental treatments. Tests of the simple effects of these treatments revealed evidence contradictory to the first part of the first hypothesis. The double confirmation procedure (CC) resulted in significantly lower error scores on acquisition tests than did the mixed procedure (PC) for the upper and lower intelligence groups. (For the middle group, the mixed procedure was superior for acquisition.) All these simple effects decreased over retention intervals.

In the case of the third hypothesis, the acquisition scores of the upper intelligence group were differentially affected by the experimental treatments. Differences were statistically significant, although they were of smaller absolute magnitude than corresponding differences in the other two intelligence groups.

Despite the uniform requirements of all subjects to write in responses to frames, training times for the treatment combinations involving confirmation (CC and PC) were significantly longer than those for the double prompting condition (PP).

It was concluded that the experimental variables were relatively weak in influencing responses, in comparison to the response control exerted by the cues in the stimulus elements in the frames, and by the concepts interrelating the frames.

TABLE OF CONTENTS

LIST OF TABLES

THE RELATIVE EFFICIENCY OF DIFFERENT COMBINATIONS OF PROMPTING AND CONFIRMATION FOR LEARNING A BOOLEAN ALGEBRA PROGRAM

SECTION I. INTRODUCTION

The study of programmed learning in the context of learning theory is gradually coming to receive the attention it deserves. As experimental studies accumulate, the weight of their evidence may clarify issues which are now in doubt, and hopefully, may lead to more satisfactory guides for the development and application of programs.

An issue which recently has been a topic for research has been the paired-associate versus the operant conditioning method of presenting the correct answer to a frame. The correct answer may be given before the subject is required to respond (prompting) or afterward (confirmation). In either case, the subject sooner or later learns to give a particular response to a particular stimulus term. In verbal learning, the simplest way to describe this is to say that an association has been formed.

Recent work by Rock (1957) and Estes (1960) suggests that associations are formed on an all-or-none basis in one trial. Estes presented a series of what he called "miniature experiments," in which a number of different kinds of stimulus objects and experimental situations were used. All supported the notion that either an association is formed or it is not, and if it is formed, only one trial is required. In these experiments, Estes found no evidence

that repetition gradually increases associative strength. On the contrary, he concluded that repeated trials in the paired-associate situation served merely to give repeated opportunities for the discontinuous formation of the learned associations between observed stimuli and responses.[1]

Estes (1960) also tested the hypothesis that after an association has once been formed, the associative strength, measured by resistance to forgetting, continues to grow as a function of reinforcement. In his experiment, following either one or two reinforcements, a sequence of two unreinforced test trials was given. Retention was measured in terms of the proportion of instances in which a correct response to a given stimulus on the first of these tests was repeated on the second test. The results failed to support the hypothesis. For paired-associate and free-verbal situations, there was no difference in the amount of retention between one and two repetitions of given material.

In a subsequent experiment by Estes, using free-verbal recall, each of two groups was given two reinforced trials. The experimental group was given a test trial between these two learning trials. Both groups were then given two successive test trials, and the amount of

---

[1] On the other hand, Dotson (1961) and Kristofferson (1961), using a different experimental technique, with paired verbal associates, found evidence that in this type of learning situation, context tended to mask increments to associative strength of previously missed pairs. In a study dealing with a simple animal learning situation, Coppock & Freund (1962), using a two-choice correction procedure, found that the probability of an errorless response in those rats which had consistently made errors was not constant, as required by the all-or-none model, but increased with trials.

retention was measured as the proportion of instances in which a correct response to a given stimulus on the first of these tests was repeated on the second test. Here Estes found a marked difference in favor of the interspersed test trial condition. Two reinforcements without an intervening test yielded only 52% retention, whereas two reinforcements with an intervening test trial yielded 78% retention.

An implication of Estes' miniature experiments seems to be that the "strength" of the association formed in one trial is dependent upon the presence of trials during learning which are like those that will be present during the test.

The question which this work suggests is: Would a combination of a prompting trial and a confirmation trial be more efficient than two trials of confirmation alone or of prompting alone, for the formation and retention of correct associations in programmed learning? It was to test this possibility and some related issues that the following experiment was performed.

SECTION II. THE EXPERIMENTAL DESIGN AND PROCEDURES

Under the experimental conditions of this study, each frame of a constructed-response program contains a stimulus element and a response element. The stimulus element is what the subject reads; the response element consists of the correct response that he is expected to make after reading the stimulus element.

The Response Modes

Under the prompting procedure, the correct response was supplied to the subject so that he could read it immediately after reading the stimulus element. In the procedure used in the experiment, the prompt was placed directly under the blank space in the frame. The subject then copied the response element which he had just read into the blank space provided.

The prompting mode is illustrated below:

A variable that increases or decreases by infinitely small amounts is called a __________ variable.
(continuous)

Under the confirmation procedure, the answer was not supplied. The subject was required to construct a response which he deemed appropriate, and to write it into the blank provided in the frame. He then turned the page to read the correct response element before going on to the next frame. Thus, in terms of the example above, the word "continuous" would not appear under the blank, but instead would be given on the back of the page.

By requiring an overt response under the prompting condition as well as the confirmation condition, the prompting-confirmation issue was not confounded with differences in overt and covert responses. The only procedural differences between the prompting and confirmation conditions consisted of the relative position in the frame at which the overt response was required (before or after the correct answer was supplied), and the longer time delay between the presentation of the stimulus element and the response element under the confirmation condition.

Combinations of Response Modes

Three combinations of prompting and confirmation were used in the study:

(1) PP (prompt-prompt). Each subject worked through all the frames of a lesson once, then immediately repeated the same lesson again. Both presentations of the lesson were given under the prompting condition, in which the subject read the stimulus element of each frame, with the correct response immediately visible under the blank space, and then wrote that response into the blank provided.

(2) PC (prompt-confirm). The subject worked through the lesson first under the prompting condition, and then immediately repeated the lesson under the confirmation condition. On the repetition, he read the stimulus element of a given frame, and constructed what he thought to be an appropriate response. Then he read the correct response on the back of the page, and found out whether his answer was correct or not.

(3) CC (confirm-confirm). The subject worked through the lesson twice, each time following the confirmation procedure described above.

The Experimental Hypotheses

The following hypotheses were to be tested by the study. The letter symbols refer to the combinations of response modes defined in the preceding section.

1. Prompting is the better condition for the initial formation of associations between the stimulus and response terms of a frame; confirmation is the better condition for strengthening associations already formed. Therefore, PC should result in fewer errors than CC when the second trial, (-C), is used as an acquisition test, and PC should result in higher scores than PP on a subsequent test of retention of associations.

2. The prompting condition does not force subjects to attend to the crucial content of the stimulus term of a frame. Hence, this condition will not be optimum for learning mediating concepts, as measured by a transfer test. Therefore, subjects who learn under either the PC or CC conditions should do better on a transfer test than those who work under the PP condition.

3. The differential effects of experimental combinations (PC, CC, or PP) will not be as large for relatively bright students as for moderate and lower intelligence levels.

The Subjects

A total of 132 high school mathematics students served as subjects. These students were divided into three intelligence levels

on the basis of their scores on the Otis (Higher Form 2A) intelligence test, administered with a 20-minute time limit. Then subjects were randomly assigned to the three experimental treatment groups (PP, PC, CC) in such a way that each group contained the same number of persons at a given intelligence level. Of the original sample, the data from 105 students were retained for the experimental analysis. The other subjects were eliminated because of absence during any of the training and testing sessions.

For the students whose data were analyzed, the median scores and their ranges for each of the three intelligence levels are shown in Table 1.

Table 1

Otis-2A Test Scores of Subjects Assigned to the Three Intelligence Levels

Intelligence Level of Group	No. of Subjects	Median Score	Range of Scores
Low	27	38	30 - 41
Moderate	36	45	42 - 48
High	42	54	49 - 64

Note.--These scores are based on tests administered with a 20-minute time limit.

Collection of the Data

Two lessons of an introductory Boolean algebra program, consisting altogether of 107 frames, were administered on successive days to the subjects. Each lesson required approximately 30 minutes.

Loose-leaf booklets, with one frame per page, were used. For the confirmation condition, answers were printed on the backs of the pages. Monitors were present during the training sessions to insure that students working under this condition followed instructions to write in a response before checking the back of the page for an answer.

Two retention intervals were used, 7 and 21 days. Two subtests were administered at the end of each of these intervals. One subtest consisted of 42 frames from the lessons, selected to represent terminal behavior. The other subtest was composed of transfer items which the students had never seen before, and which required generalization of concepts. This subtest contained 33 items. The same frames and items were used at both retention intervals, but were arranged in a different random order for the second interval.

Types of Scores

Several different scores were used, and were analyzed separately:

(1) The error score was defined as the number of frames in which the wrong response had been written. These scores were obtained for all acquisition trials under the confirmation condition, and on all retention interval tests. Each subject's error score for the 33 transfer item subtests was multiplied by 42/33 to simplify the analyses of variance involving both repeated items and transfer items.

(2) The proportion-retained retention score was defined as the proportion of frames correct on the first retention test that were also correct on the second retention test.

(3) The <u>raw retention score</u> was the number of wrong responses on either repeated-frame or transfer subtests given at the 7-and 21-day retention intervals.

(4) The <u>time score</u> was the amount of time, to the nearest minute, taken during the learning session.

SECTION III. RESULTS

The statistical analyses are summarized in Tables 2 through 8.

Analysis of the Error Scores

Table 2 presents mean error scores for the PC and CC experimental groups, the two groups in which acquisition scores could be obtained. For the PC combination, Trial 1 responses were prompted, and consequently there were no opportunities to make errors. Overall error scores on Trial 2 were, however, about as high as those for the corresponding trial for the CC combination. The low and high groups had low error scores on Trial 2 for the CC combination; the middle group had low error scores under the PC condition. In statistical terms, this suggests a significant interaction between intelligence levels and treatment groups. In the discussion of the analysis of variance of these data, which follows, this suggested interactive effect is confirmed.

The analysis of variance of these acquisition error data (Trial 2) for the PC and CC groups, and of the error scores for the repeated-item subtests for these two groups, is summarized in Table 3. In this analysis, Trial 2 data were analyzed with retention test data. These error scores are based, then, on one immediate and two delayed (7 and 21 days) tests.

Differences among retention levels are seen to be significant at the .01 level, as are the differences among intelligence levels.

Table 2

Means and Standard Deviations of Acquisition Error Scores Obtained by CC and PC Groups Arranged by Intelligence Levels

Treatment Group	Intelligence Level	N	Trial 1[a] Mean	Trial 1[a] SD	Trial 2[b] Mean	Trial 2[b] SD
CC	Low	9	11.00	7.79	6.78	6.13
	Average	12	12.58	5.63	8.42	2.72
	High	14	7.21	4.33	2.21	2.51
	All levels	35	10.03	6.30	5.51	4.73
PC	Low	9	--	--	9.56	4.00
	Average	12	--	--	3.25	2.83
	High	14	--	--	4.93	3.94
	All levels	35	--	--	5.54	4.38

[a]On Trial 1 for the PC combination, there were no opportunities to make errors.

[b]All gains between Trial 1 and 2 were statistically significant at the .05 level for the Low and Average groups, and at the .01 level for the High groups or for the combined subgroups of either treatment group.

On the other hand, the main effect for treatment groups (CC vs. PC) is not significant. Since intelligence levels and treatments interact significantly, the nonsignificance of this treatment effect apparently can be attributed to the canceling effects of various intelligence levels within a given treatment category.

Table 3

Analysis of Variance of Acquisition and Retention Error Scores for Repeated-Frame Subtests for CC and PC Groups

Source of Variation	df	Sum of Squares	Mean Square	F
Between Subjects	69	6,567.60		
Intelligence levels	2	2,573.75	1,286.87	24.225**
Treatments (CC vs. PC)	1	43.89	43.89	0.827
Intelligence levels X treatments	2	550.20	275.09	5.178**
Error (b)	64	3,399.76	53.121	
Within Subjects	140	6,194.00		
Retention intervals	2	3,378.94	1,689.47	89.456**
Retention intervals X intelligence levels	4	358.59	89.64	4.747*
Retention intervals X treatments	2	29.34	14.67	.777
Retention intervals X intelligence levels X treatments	4	9.76	2.44	.129
Error (w)	128	2,417.37	18.886	

*Significant at .05 level.

**Significant at .01 level.

There was also a significant interaction between intelligence levels and retention intervals, due primarily to the differences among immediate and 7-day retention interval means for the three intelligence levels.

Table 4 summarizes mean error scores on retention subtests (7-and 21-day intervals, repeated and transfer items) for all three treatment groups and the three intelligence levels. In general, the subjects did better on the repeated-frame subtests than on the transfer subtests. The differences between repeated-frame and transfer subtest scores were statistically significant at the .01 level. Differences between the 7-day and 21-day retention intervals were not significant. As mentioned in connection with the analysis of Table 3, most forgetting occurred in the first seven days. The analysis of variance of the 7- and 21-day retention test data in Table 4 is presented in Table 5.

Analysis of the Proportion Retained Retention Scores

A measure of retention available from repeated tests is the number of items correctly answered on the second test that were also answered correctly on the first. This measure has been used in paired-associate studies (Estes, 1960). There the assumption was that associations are formed on an all-or-none basis, and hence that the proportions of correct responses following incorrect responses on repeated tests should be no better than chance.

This measure was computed for the repeated-frame subtest, for the 7- and 21-day retention intervals. These proportions are summarized in Table 6. The use of this criterion of retention partials

Table 4

Summary of Mean Error Scores on Retention Subtests Obtained by Three Treatment Groups at Three Intelligence Levels

Intelligence Level	N	Retention Intervals: Repeated Frames 7 days	Repeated Frames 21 days	Transfer Items[a] 7 days	Transfer Items[a] 21 days
(Confirmation-Confirmation group)					
Low	9	18.11	18.89	23.90	24.77
Moderate	12	18.33	19.75	23.98	24.18
High	14	9.21	8.00	14.46	12.00
All levels	35	14.63	14.83	20.15	19.46
(Prompting-Confirmation group)					
Low	9	20.00	19.33	25.33	28.28
Moderate	12	13.33	13.75	19.95	18.98
High	14	9.86	8.36	15.27	13.27
All levels	35	13.66	13.03	19.46	19.09
(Prompting-Prompting group)					
Low	9	21.44	21.78	29.29	27.73
Moderate	12	15.58	14.50	19.72	19.62
High	35	9.50	9.79	15.54	13.10
All levels	35	14.66	14.49	20.51	19.10

[a]Weighted by a factor of 1.273 to equate the 33 transfer items with the 42 repeated frames.

Table 5

Analysis of Variance of Retention Error Scores for Three Treatment Groups at Three Intelligence Levels

(Based on scores obtained on tests at 7- and 21-day intervals)

Source of Variation	*df*	Sum of Squares	Mean Square	*F*
Between Subjects	104	23,722.24		
Intelligence levels	2	9,540.07	4,770.03	34.679**
Treatments (PP, PC, CC)	2	79.13	39.56	.288
Intelligence levels X treatments	4	897.88	224.47	1.632
Error (b)	96	13,205.16	137.55	
Within Subjects	315	6,504.89		
Retention intervals	1	27.67	27.670	2.534
Retention intervals X intelligence levels	2	79.34	39.670	3.653*
Retention intervals X treatments	2	5.25	2.675	.240
Retention intervals X intelligence levels X treatments	4	28.31	7.053	.646
$Error_1$ (w)	96	1,048.16	10.917	
Item type (Repeat vs. transfer)	1	3,078.02	3,078.020	179.518**
Item type X intelligence levels	2	55.95	27.975	1.632
Item type X treatments	2	14.55	7.275	.424
Item type X intelligence levels X treatments	4	7.17	1.793	.105
$Error_2$ (w)	96	1,646.00	17.146	
Retention intervals X item type	1	10.31	10.310	2.292
Retention intervals X item type X intelligence levels	2	18.12	9.060	2.014
Retention intervals X item type X treatments	2	10.72	5.360	1.192
Retention intervals X item type X intelligence levels X treatments	4	43.60	10.900	2.423
$Error_3$ (w)	96	431.82	4.498	
Total	419	30,227.13		

*Significant at .05 level.

**Significant at .01 level.

out the effects of differences in the amount originally learned. Hence, Table 6 indicates that rate of forgetting is correlated with intelligence. However, there is no evidence in Table 6 that the different experimental conditions affected strength of associations differentially, except for the moderate, or middle, intelligence group. Their retention scores were slightly worse under the CC condition than under the other two conditions.

Table 6

Proportions of Same Frames Correct on Both First (7-day) and Second (21-day) Retention Tests

Intelligence Level	Treatment Group		
	CC	PC	PP
Low	.749	.754	.751
Moderate	.753	.836	.829
High	.928	.919	.884
All levels	.822	.848	.831

Analysis of the Time Scores

There is some practical interest in the times required for acquisition under different procedures. It is obvious that where write-in responses are compared with covert or implicit responding, the former will be slower. However, in this study, write-in responses were always required, no matter what combination of prompting

and confirmation were involved. Nevertheless, as can be seen in Table 7, the PP condition yielded lower training time means. This reflects the time required for the subject to try to think of the correct answer under the conditions involving confirmation.

Table 7

Means and Standard Deviations of Training Times (Minutes) for Three Treatment Groups at Three Intelligence Levels

Intelligence Level	Treatment Group					
	CC		PC		PP	
	Mean	SD	Mean	SD	Mean	SD
Low	63.1	14.3	63.1	7.9	50.3	14.3
Moderate	60.9	9.6	58.7	10.6	51.7	10.5
High	59.2	8.2	52.3	6.8	43.9	8.6
All levels	60.8	10.9	57.3	9.6	48.2	10.9

The analysis of variance of these time scores, shown in Table 8, indicates that the time differences between experimental groups are significant at the .05 level.

Table 8

Analysis of Variance of Training Times for Three Treatment Groups at Three Intelligence Levels

Source of Variation	Sum of Squares	*df*	Mean Square	*F*
Intelligence levels	2,941.257	2	1,470.629	13.817**
Treatments	965.080	2	482.540	4.533*
Intelligence levels X treatments	237.207	4	59.302	.557
Error (within cells)	10,218.171	96	106.439	
Total	14,361.715	104		

*Significant at .05 level.

**Significant at .01 level.

SECTION IV. DISCUSSION

The study found inconclusive support for the hypothesis that the PC combination is superior to CC for acquisition and to PP for retention. Treatments were not significant as a main effect in the analyses of variance (Tables 3 and 4). During acquisition, subjects in the high and low intelligence levels made fewer errors under the CC than under the PC combinations. However, those in the middle intelligence level made fewer errors under the PC than the CC combination during acquisition (Table 2). While some of these simple effects were statistically significant, they diminished over the 7- and 21-day retention intervals. The brightest students did learn more and did retain it better than the others, which led to a significant interaction between intelligence levels and retention intervals (Tables 3 and 5). The data for the transfer-item subtests are negative. The treatment combinations had no statistically significant differential effects on transfer of concepts to new problems (Tables 4 and 5).

These results are in general agreement with other studies of prompting and confirmation summarized by Silberman (1961). When these variables are studied in the context of programmed instruction, either no differences are found, or differences are small and of little practical significance.

Results of paired-associate and simple rote learning studies concerned with these variables (Cook & Kendler, 1956; Angell &

Lumsdaine, 1960; Stolurow, 1961) are usually more definitive. Paired-associate studies have found strong evidence that prompting is a superior condition for acquisition of associations. Stolurow found that prompting groups did better than confirmation groups on both recall and recognition tests, but that amount of overlearning also influenced the obtained results. Angell and Lumsdaine, using Cook's stimulus materials, found prompting in only three trials out of four superior to either prompting or confirmation alone on every trial.

Although there is a temptation to assume that the paired-associate model fits programmed instruction of verbal material, the two learning situations clearly are different in important respects. The stimulus terms of paired associates are carefully constructed not to have prior associational value for their response terms, and interitem associations are similarly minimized. The opposite is true of programmed instruction. Frames are carefully constructed to make correct responses obvious; highly overlearned verbal symbols are used; and mediating concepts interrelate the successive frames. Under these conditions, it is likely that prompting and confirmation are relatively weak variables in comparison to the response control exerted by the cues in the stimulus elements and by the concepts interrelating the frames.

REFERENCES

Angell, D., & Lumsdaine, A. A. Prompted tests, prompted trials, versus prompted trials alone in paired-associate learning. Pittsburgh: American Institute for Research, October 1960.

Cook, J. O., & Kendler, T. S. A theoretical model to explain some paired-associate learning data. In G. Finch & F. Cameron (Eds.), Symposium on Air Force human engineering, personnel, and training research. Washington, D. C.: National Academy of Sciences - National Research Council, 1956. (Publ. 455) Pp. 90-98.

Coppock, H. W., & Freund, J. E. All-or-none versus incremental learning of errorless shock escapes by the rat. Science, 1962, 135, 318-319.

Dotson, J. M. Unpublished M.A. thesis, Univer. of Cincinnati, 1961.

Estes, W. K. Learning theory and the new "mental chemistry." Psychol. Rev., 1960, 67, 207-223.

Kristofferson, A. B. Repetition and paired associates learning. Science, 1961, 134, 2036-2037.

Rock, I. The role of repetition in associative learning. Amer. J. Psychol., 1957, 70, 186-193.

Silberman, H. F. Characteristics of some recent studies of instructional methods. Paper read at Conference on Application of Digital Computers to Automated Instruction, cosponsored by the Office of Naval Research, and System Development Corp., Washington, D. C., October 1961.

Stolurow, L. M. Prompting vs. confirmation sequences and overlearning in the automated teaching of sight vocabulary. Paper read at XIV Int. Congr. of Appl. Psychol., Christiansborg Castle, Copenhagen, Denmark, August 1961.